

# Retention capacity of random surfaces

Craig L. Knecht<sup>1,\*</sup> Walter Trump<sup>2,†</sup> Daniel ben-Avraham<sup>3,‡</sup> and Robert M. Ziff<sup>4,§</sup>

<sup>1</sup> 691 Harris Lane, Gallatin TN 37066 USA

<sup>2</sup> Department of Physics, Gymnasium Stein, 90547 Stein, Germany

<sup>3</sup> Department of Physics, Clarkson University, Potsdam NY 13699-5820 USA and

<sup>4</sup> Michigan Center for Theoretical Physics and Department of Chemical Engineering, University of Michigan, Ann Arbor MI 48109-2136 USA

(Dated: February 1, 2012)

We introduce a “water retention” model for liquids captured on a random surface with open boundaries, and investigate the model for both continuous and discrete surface heights  $0, 1, \dots, n-1$  on a square lattice with a square boundary. The model is found to have several intriguing features, including a nonmonotonic dependence of the retention on the number of levels: for many  $n$ , the retention is counterintuitively greater than that of an  $n+1$ -level system. The behavior is explained using percolation theory, by mapping it to a 2-level system with variable probability. Results in one dimension are also found.

PACS numbers: 64.60.ah, 64.60.De, 05.50.+q

Consider a bounded horizontal random surface with a landscape of varying height, as shown in Fig. 1. A liquid such as water is dripped over the surface and is allowed to drain out all of the boundaries. Internal sites in valleys capture the water and create ponds, and eventually all the ponds fill up to their maximum height. We are interested in finding the total amount of water retained in the system when the maximum heights are reached. Physically, this problem is related to coatings on a random surface and the properties of landscapes and watersheds. Theoretically, it is related to the topology of random surfaces [1, 2] and to invasion percolation (IP), but with some interesting new features.

We study this problem on a regular square lattice with random heights assigned to each site. The systems are square of size  $L \times L$  with draining boundaries on all four sides. Extensive simulations were performed with uniformly distributed discrete heights  $0, 1, 2, \dots, n-1$  for values of  $n$  ranging from 2 to 100, and also for a continuum of heights  $0, 1$ . We also studied a 2-level system with variable occupation probabilities of the 2 heights. The simulation method we used is a form of IP in which we effectively reversed the flow and flooded the system from the outside with higher water levels, and recorded the level of the water in a pond when it was first flooded. The retention is the difference between that level and the height of the terrain below the pond.

Fig. 2 shows the average retention  $R_n^{(L)}$  on  $n$ -level systems, for  $n = 2, \dots, 8$ , as a function of  $L$ . Here we assume that all terrain heights occur with equal probability. As expected, the retention grows as  $L^2$  for large  $L$ , and generally grows with  $n$ , as more levels create deeper ponds. However, we found deviations to this expected behavior. As seen in Fig. 2, there is a crossover in the curves for  $n = 2$  and  $3$ : for small  $L$ ,  $R_2^{(L)} < R_3^{(L)}$ , but for  $L > 51$ ,  $R_2^{(L)} > R_3^{(L)}$ . This is in spite of the fact that 3-level systems can have ponds of water of depth 2, while 2-level

systems can only have ponds of depth 1. Further study shows additional crossings between levels  $n$  and  $n+1$  at seemingly random  $n$ 's, and at larger values of  $L$  (Table I).

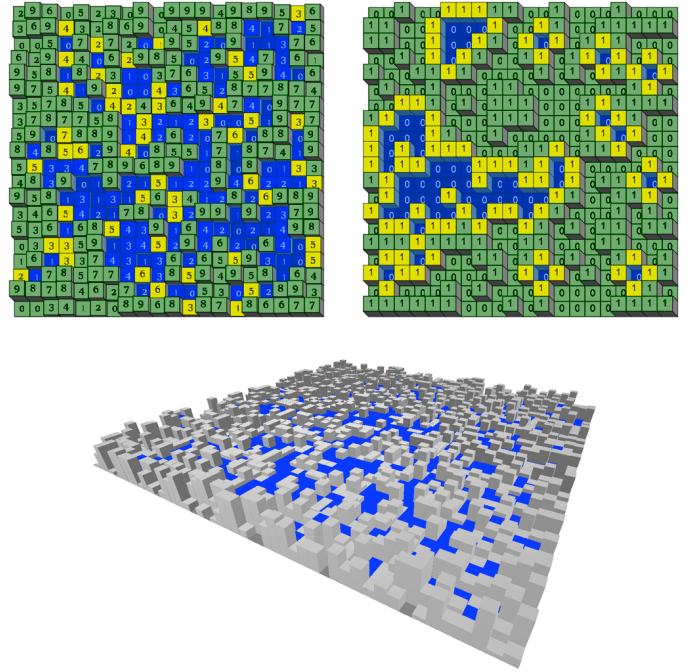


FIG. 1. (color online) Examples of water retention on square systems with an equal distribution of terrain heights. Top left: heights 0 through 9; Top right: heights 0 and 1. Green (gray with black height numbers) are dry and above water, blue (gray with white numbers) are filled with water, and yellow (light shade) are spillover sites of the height of an adjacent pond (and only shown if neighboring a wet site). Bottom: Perspective view of a water-filled 10-level system, with dry sites shown as grey (lighter shading on top).

In this Letter we explain some of the puzzling features of this model, though many questions remain. Some re-

lated issues, especially involving multi-level nonuniform systems, are discussed in [3].

To analyze the multi-level discrete model, we make a decomposition of  $R_n^{(L)}$  in terms of the retention in a 2-level system with varying  $p$ ,  $R_2^{(L)}(p)$ , where  $p = \text{Prob}(0)$  is the probability or fraction of sites with terrain height 0 in the 2-level system:

$$R_n^{(L)} = \sum_{i=1}^{n-1} R_2^{(L)}\left(\frac{i}{n}\right). \quad (1)$$

The  $i = 1$  term represents the amount of water retained up to just the first level, for which all sites of terrain height 1 or higher can be considered as level 1. The net fraction of 0-height sites is  $1/n$ . The  $i = 2$  term represents the total amount of water at just the second level; here we collapse the first 2 levels into the new level 0 (fraction  $2/n$ ), and the rest of the sites can be considered as level 1. Likewise, the remaining terms follow.

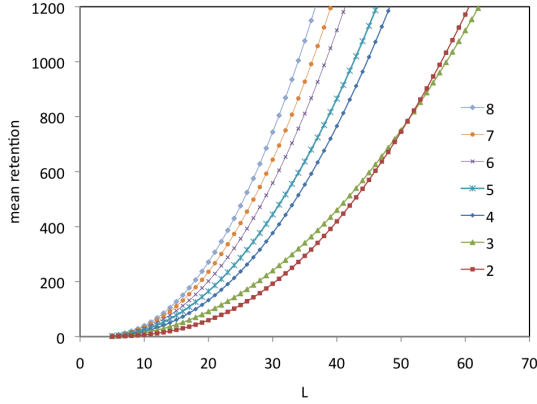


FIG. 2. (color online) Retention  $R_n^{(L)}$  vs.  $L$  for  $n = 2, 3, \dots, 8$  with a uniform distribution of levels, showing the crossing of the curves for  $n = 2$  and  $3$  at  $L \approx 51$ . Additional crossings are found for larger  $L$  (see Table I). For small  $L$  there is an ordering of the retention with  $n$ , which is violated for larger  $L$ .

It is also possible to show that  $(1/i)R_2^{(L)}(i/n)$  is equal to the number of sites with retention  $i$ . Thus, the total number of wet sites is

$$W_n^{(L)} = \sum_{i=1}^{n-1} \frac{1}{i} R_2^{(L)}\left(\frac{i}{n}\right). \quad (2)$$

It is therefore necessary to just know the behavior of the 2-level system with a varying  $p$  in order to predict the behavior of all  $n$ -level systems (including those with nonuniform level distributions). We have carried out simulation of  $R_2^{(L)}(p)$  for  $p = 0.01, 0.02, \dots, 0.99$  for various  $L$ , and the results are shown in Fig. 3. For small  $L$ , the curve is rounded and peaked close to  $P(0) = 1/2$ , but for larger  $L$  it approaches a ramp of slope 1, up to the

value of  $p = p_c = 0.592\,746$  (the site percolation threshold), after which it drops off precipitously and rapidly approaches 0. This behavior is best understood from the limit  $L \rightarrow \infty$ . Define  $r_2(p) = \lim_{L \rightarrow \infty} R_2^{(L)}(p)/L^2$ . In an infinite system, all finite clusters of 0-sites retain water and the infinite cluster alone drains off. Thus, the total retention is

$$r_2(p) = p - P_\infty(p), \quad (3)$$

where  $P_\infty(p)$  is the fraction of sites belonging to the infinite cluster. Very close to (and above)  $p_c$ ,  $P_\infty \sim a(p - p_c)^\beta + \dots$  where  $a$  is a constant and  $\beta = 5/36$  [4]. Because  $P_\infty$  rises quickly as  $p$  increases beyond  $p_c$ , we get the quick drop to zero in  $r_2(p)$ . For finite  $L$ , finite clusters at the boundaries drain as well, yielding the observed finite-size effects.

Exactly at  $p_c$ , the drainage area is fractal, yet the retained water is still proportional to  $L^2$  for large  $L$  with corrections proportional to  $L^{d_f}$  where  $d_f = 91/48$  is the fractal dimension. We verified that at  $p_c$ , the size distribution of the draining clusters (boundary clusters in percolation) satisfies  $n'_s \sim s^{-\tau'}$  with  $\tau' = 1/d_f - 1 = 139/91 \approx 1.527$  as predicted in [5]. Our measurement of  $\tau' = 1.5256 \pm 0.003$  confirms this prediction to about 20 times the precision given in [5].

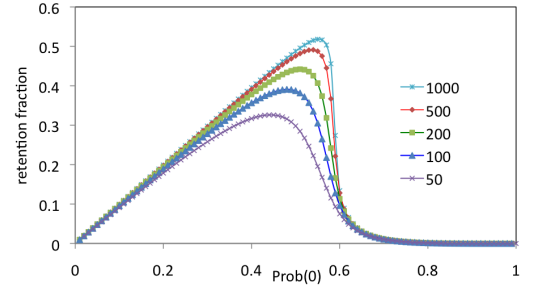


FIG. 3. (color online) Retention fraction  $R_2^{(L)}(p)/L^2$  in a 2-level system with varying  $p = \text{Prob}(0)$ , for  $L$  ranging from 50 to 1000, averaged over 10 000 to 1 000 000 samples for each point.

As a good approximation for large  $L$ , we can ignore the small contribution to  $R_2^{(L)}(p)$  for  $p > p_c$  and approximate  $R_2^{(L)}(p) = pL^2$  for  $p < p_c$ . Then, from (1), we find the following formula for the  $n$ -level retention in the large- $L$  limit:

$$R_n^{(L)}/L^2 \approx \sum_{i=1}^{n^*} \frac{i}{n} = \frac{n^*(n^* + 1)}{2n} \quad (4)$$

where  $n^*$  is the largest integer such that  $n^*/n$  is less than  $p_c$ . Thus, we have  $R_2^{(L)} \sim (1/2)L^2$ , and  $R_3^{(L)} \sim (1/3)L^2$ , which indeed gives  $R_2^{(L)} > R_3^{(L)}$  for large  $L$ . This result can be explained simply by the fact that for the 2-level system, roughly half the sites are 0's and filled with water, while for the 3-level system, only  $1/3$  of the sites

are 0's; very few ponds are filled to a level of 2 because those sites correspond to clusters above the percolation threshold.

To explain the crossing, we must also explain why the curves for  $R_n^{(L)}$  are ordered  $R_2^{(L)} < R_3^{(L)} < R_4^{(L)} \dots$  for small  $L$ . This can be understood qualitatively from the behavior of  $R_2^{(L)}(p)$  for small  $L$  as in Fig. 3: because those curves are smooth, equation (1) will be a gradual, increasing function of  $n$ . To be more rigorous, we consider the smallest system possible: a  $3 \times 3$  system, which has only one site that can hold water (the center site), and only four sites that can block it, as the corner sites are irrelevant. A direct calculation yields

$$R_n^{(3)} = (n^2 - 1)(3n^2 - 2)/(60n^3), \quad (5)$$

which is a monotonically increasing function of  $n$ . (Details of the derivation will be given in a future paper.) Because the ordering is verified for  $L = 3$  but not for large  $L$ , crossing must necessarily occur for some  $L$ .

The curves that are “out of order” and cross are those in which  $r_n = n^*(n^* + 1)/(2n)$  is greater than  $r_{n+1}$ , by (4). This occurs when the fractional part of  $p_c n$  is between 0 and  $1 - p_c \approx 0.407$ . The crossing curves  $(n, n + 1)$  are at  $(2, 3)$ ,  $(4, 5)$ ,  $(7, 8)$ ,  $(9, 10)$ ,  $(12, 13)$ ,  $(14, 15)$ ,  $(17, 18)$ , etc. We have verified the first six crossings as shown in Table I. For  $n > 30$ , the simple analysis based upon (4) evidently breaks down as contributions from  $R_2(p)$  for  $p > p_c$  become important, and the crossings are predicted to become less frequent, though we have not measured them directly.

TABLE I. Crossing points where  $R_n^{(L^*)} = R_{n+1}^{(L^*)}$ , extrapolated to non-integer  $L^*$ .

$n$ and $(n + 1)$	$L^*$	$R_n^{(L^*)}$
2 and 3	51.12	790
4 and 5	198.1	26 000
7 and 8	440.3	246 300
9 and 10	559.1	502 000
12 and 13	1390.6	4 288 500
14 and 15	1016.3	2 607 000

In the limit that the number of levels becomes infinite, the discrete system goes over to the continuum one. Now, as in traditional IP [6], the fluid flows over the lowest barrier site on the perimeter of a pond. For a continuum bond IP system, the “raining” IP problem has recently been considered in [7, 8], and the pond-size distribution, away from the boundaries, was investigated.

Here, considering the continuum site system, we find that water rises to an average height of  $\bar{h} \approx 0.6039$  (averaged over wet sites only, for  $L \rightarrow \infty$ ), which is slightly above  $p_c$ . The large ponds have a water level that is slightly below  $p_c$ , because higher levels produce large percolation clusters that would run into the boundary.

There are also small ponds with higher levels, corresponding to clusters above the threshold; these allow the average water level to be above  $p_c$ . Fig. 4 shows the water level of sites when first flooded as a function of the number of sites flooded, showing the small contribution of the ponds of high level.

In fact, taking the continuum limit of (1) we can calculate the total retention  $r$  per site in the continuum system directly by integrating the curve of  $r_2(p)$ ,  $r = \int_0^1 r_2(p) dp$ . The triangular part below  $p_c$  gives  $p_c^2/2$  exactly, and the tail above  $p_c$  gives a small correction. The tail's area extrapolates to 0.0063 for large  $L$ , and predicts  $r = p_c^2/2 + 0.0063 = 0.1820$ , which we verified directly to  $\pm 0.0002$  by measuring the retention for systems of up to  $L = 12000$  and extrapolating to  $L = \infty$ . The retention per site  $r$  is equal to  $\langle wh \rangle / 2 \approx w \bar{h} / 2 \approx \bar{h}^2 / 2$  where  $w = W^{(L)} / L^2$ . Note,  $w = \int_0^1 (r_2(p)/p) dp$  follows from (2) in the continuum limit, and we find  $w = p_c + 0.0100 = 0.6028$  in agreement with direct measurement (see (7));  $\bar{h} = 2r/w = 0.6039$  was also independently measured.

In fact, applying (3) to Eqs. (2) and (1) we see that  $w = 1 - \mu_{-1}$ ,  $r = 1/2 - \mu_0$ , and more generally, we find the moments of the retention as

$$\langle r^q \rangle = \lim_{L \rightarrow \infty} \frac{\langle (R^{(L)})^q \rangle}{L^2} = \frac{1}{q+1} - \mu_{q-1}, \quad (6)$$

where  $\mu_q = \int_0^1 x^q P_\infty(x) dx$  is the  $q$ -th moment of  $P_\infty$ . Thus, we have found that the moments of  $P_\infty$  assume a specific physical interpretation in the context of the retention problem.

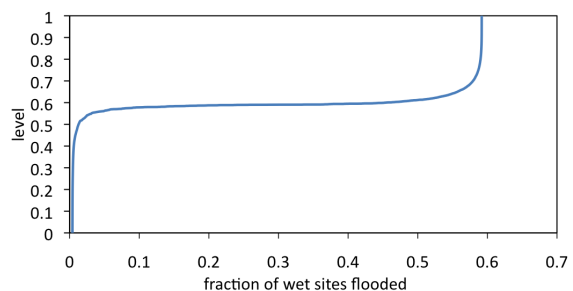


FIG. 4. Level when sites are first flooded vs. the fraction of sites flooded on a  $1000 \times 1000$  continuum system.

The asymptotic behavior of  $W^{(L)}$  and  $R^{(L)}$  in the continuum system is found to be

$$\begin{aligned} W^{(L)} &\sim 0.6028L^2 - 2.4L^{1.25} \\ R^{(L)} &\sim 0.1820L^2 - 1.27L^{1.25} \end{aligned} \quad (7)$$

where the second terms reflect the effects of the drainage sites and ponds of lower water level near the boundary. At each level in a discrete system, the drainage area is just all the clusters touching the boundary, which extends into the system a distance of the correlation length

$\xi \sim |p - p_c|^{-\nu}$ . Integrating this over  $p$  up to  $p_c - cL^{-1/\nu}$  and multiplying by the perimeter  $4L$  gives a depletion zone  $\propto L^{2-1/\nu} = L^{5/4}$ . The value of the exponent 1.25 was verified numerically to  $\pm 0.05$ . Recently, it has been shown that watersheds, bounded by the “continental divide” between drainage regions, have a similar fractal dimension  $d_f \approx 1.22$  [9]. It appears that these two problems, however, are different, despite the similarity of exponents.

The assignment of a terrain height for each site using a probability corresponds to a grand canonical type of description. One can also distribute the levels canonically, with exactly  $1/n$  of them of each height. We carried out simulations using this ensemble and found only small differences. For  $L = 3, 4$ , and  $5$ , we also carried out an exact enumeration of all canonical states. For small  $L$ ,  $\Delta = R_n^{(L)}[\text{canonical}] - R_n^{(L)}[\text{grandcanonical}] > 0$ , and for larger  $L$ ,  $\Delta$  decreases. For  $n = 2$ ,  $\Delta$  appears to approach 0, while for  $n = 3, 4, 5$ ,  $\Delta$  appears to approach a negative constant for large  $L$ . Because the value of the retention itself grows as  $L^2$ , the relative difference  $\Delta/R_n^{(L)}$  is very small. We verified that changing to the canonical ensemble does not affect the crossing behavior of the  $R_n^{(L)}$  curves.

We studied the distribution  $n_s$  of ponds of  $s$  sites and verified that the system self-organizes to the percolation critical point with  $n_s \sim s^{-\tau}$  and  $\tau = 187/91$ . Unlike standard percolation, we cannot write exact formulas for any  $n_s$ —not even for  $n_1$ . However, we can make an estimate for  $n_1$  as follows: The probability that a site is in a pond of size 1, of water height between  $x$  and  $x + dx$ , is given approximately by

$$P_1(x)dx = 4(1-x)^3x[1-(1-x)^3]dx \quad (8)$$

where the factor of  $4(1-x)^3$  is the probability that 3 of the neighbors are of higher terrain height (4 possibilities),  $x$  is the probability that the site itself is of terrain height less than or equal to  $x$ , and  $[1-(1-x)^3]$  is the probability that the spillway site has at least one neighbor lower than  $x$ , so the spillway can drain at least to the next sites. This gives  $n_1 \approx \int_{x_0}^1 P_1(x)dx = 0.01624$  where  $x_0 = \bar{h} = 0.6039$  is the average water height surrounding the cluster. This compares to a measurement of  $n_1 = 0.01595$ . Likewise, the average water height of the ponds of size 1,  $\bar{h}_1 = \int_{x_0}^1 xP_1(x)dx / \int_{x_0}^1 P_1(x)dx = 0.6904$ , is close to the extrapolated measured value 0.6887.

We also studied the model in one dimension (1D), where there are no crossings, however exact results for all quantities can be found. Consider, for example, the semi-infinite line (sites  $1, 2, 3, \dots$ ) so that water can spill only through the left edge, and assume a uniform distribution of barriers in  $[0, 1]$ . As we look at the 1D ponds, starting from the left edge, the water level keeps rising the farther we venture into the line. In fact, each pond begins when a record-height barrier is encountered, and

ends when the next, yet higher barrier, is met.

The probability that the barrier at site  $k+1$  is taller than all the  $k$  preceding barriers is  $\int_0^1 x^k dx = 1/(k+1)$ . This is also the probability that a pond starts (or ends) at site  $k+1$ . Because the barriers demarcating the ponds occur with probability  $1/k$  at site  $k$ , it follows that the typical size of ponds,  $k$  sites away from the edge, is  $k$ . The ponds grow linearly with their distance from the edge.

Next consider  $p_s(k)$ , the probability that a 1D pond of size  $s$  is  $k$  sites away from the edge, in sites  $k+1, k+2, \dots, k+s$ . For that pond to have water level  $x$ , the first  $k-1$  sites must have barriers lower than  $x$  (with probability  $x^{k-1}$ ), as do sites  $k+1, k+2, \dots, k+s$  (probability  $x^s$ ). Site  $k$  contains a barrier of height  $x$  (probability  $dx$ ), and site  $k+s+1$  contains a barrier of height  $y > x$  (probability  $1-x$ ). Thus the probability for a pond of level  $x$  is  $x^{k-1+s}(1-x)dx$ . Integrating over  $x$ , we obtain the required probability:

$$p_s(k) = \int_0^1 x^{k-1+s}(1-x)dx = \frac{1}{(s+k)(s+k+1)}. \quad (9)$$

Note that  $\sum_{s=1}^{\infty} p_s(k) = 1/(k+1)$ , consistent with our previous result, and that the moments of  $p_s(k)$  diverge, which is why we estimated the typical pond size instead.

Similarly, the probability for having  $k$  draining sites at the edge is

$$p_{\text{drain}}(k) = \frac{1}{(k-1)!} \int_0^1 x^k dx = \frac{k}{(k+1)!}. \quad (10)$$

These results illuminate the analogous quantities in 2D, where however no exact results could be found.

In this Letter we have only touched upon the questions that one may ask about the retention model. There are many more questions that are unsolved, including the exact results for the size distribution of the clusters, the average retention as a function of the distance from an edge, the behavior on other lattices, on systems with different boundary shapes, in higher dimensions, and systems with a tilt. We believe it is an interesting model that warrants much further study.

We mention finally that the water retention problem was previously studied in the context of surfaces created by magic squares [10]. The application to random surfaces is an example of the deeper connections of this problem.

*Acknowledgments:* The authors acknowledge correspondence with Neal Madras, Gareth McCaughan, and Seung Ki Baek. Assistance from Spencer Snow and Joe Scherping is also noted.

---

\* craigknecht03@gmail.com

† w@trump.de

<sup>†</sup> qd00@clarkson.edu

<sup>§</sup> rziff@umich.edu

- [1] S. N. Majumdar and O. C. Martin, Statistics of the number of minima in a random energy landscape, Phys. Rev. E **74**, 061112 (2006).
- [2] S. Carmi, P. L. Krapivsky, and D. ben-Avraham, Partition of networks into basins of attraction, Phys. Rev. E **78**, 066111 (2008).
- [3] S. K. Baek and B. J. Kim, arXiv 1111.0425.
- [4] D. Stauffer and A. Aharony, Introduction to Percolation Theory (Taylor and Francis, Philadelphia, 1994).
- [5] T. Lookman and K. De'Bell, Surface fractal dimension for percolation clusters: A Monte Carlo study, Phys. Rev. B **46**, 5721 (1992).
- [6] D. Wilkinson and J. F. Willemsen, Invasion percolation: a new form of percolation theory. J. Phys. A **16** 336 (1983).
- [7] J. van den Berg, A. A. Jarai, and B. Vágvolgyi, The size of a pond in 2D invasion percolation, Electronic Comm. Probab. **12**, 411 (2007).
- [8] M. Damron, A. Sapozhnikov and B. Vágvolgyi, Relations between invasion percolation and critical percolation in two dimensions. Ann. Probab. **37**, 2297 (2009).
- [9] E. Fehr, D. Kadau, N. A. M. Araújo, J. S. Andrade, and H. J. Herrmann, Scaling relations for watersheds, Phys. Rev. E **84**, 036116 (2011).
- [10] C. L. Knecht,  
http://www.knechtmagicsquare.paulscomputing.com/

---

Note added: This paper was published in Physical Review Letters **108**, 045703 (January 25, 2012). Some additional material is given in the web page

[http://en.wikipedia.org/wiki/Water\\_retention\\_on\\_mathematical\\_surfaces](http://en.wikipedia.org/wiki/Water_retention_on_mathematical_surfaces),

where retention on magic squares is also discussed. Measurements of the crossing points  $R_n^{(L^*)} = R_{n+1}^{(L^*)}$  for several larger values of  $n$  have been found and will be given in a future publication.

---